



Key Curriculum Press
Innovators in Mathematics Education

Certificate of Completion

is granted to

MICHAEL CHAO

for participation in the
Teaching Statistics with Fathom Online Course in
Emeryville, CA

This professional development seminar consisted of thirty hours of
online instruction, as well as additional participant study.

Karen Coe, President

November 3-December 22, 2008
Emeryville, CA

TSF NOV 3 SEC 1: Assignment: Probability Model Project - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines 10day weather - Goo... Customize Links Radio Station Guide Windows Media Windows http://www.culture.s...

TSF NOV 3 SEC 1: Assignment: Pro... iGoogle

http://pdonline.keypress.com/mod/assignment/view.php?id=30768

Due date: Monday, November 24, 2008, 12:55 PM

Feedback from the Moderator



Corey Andreasen

Sunday, November 30, 2008, 01:44 PM

Grade: 4 / 4

Michael,

Nice job on the Week 3 simulation. I see your concern about too many levels of depth being hard to follow. That can be an issue when working with students. One solution would be to collect more measures (maybe several thousand) in your first measures collection, and check the proportion that were 5 (maybe in a summary table).

Sampling without replacement works well here, but you should also be aware of the Scramble Attribute Values feature, which is the way I would have approached this problem. (That doesn't mean it's better than your approach, but it's the way I'd think about it.) But the Scramble Attribute Values is a terrific way to simulate things like randomized experiments under the null hypothesis, so you should be aware of it.

Like I said, Nice work!

Corey

ChaoM WK3 Proj.frm

start iGoogle - Windows In... The New York Times - ... TSF NOV 3 SEC 1: As... 10:32 PM






This document is for problem E21, the backpack problem:

A group of 5 friends have their backpacks that all look alike. They toss their backpacks on the ground and later pick up a backpack at random. Estimate the probability that everyone gets his or her own backpack.

I'm pretty sure that the theoretical probability answer is $1/5!$ or $1/120$.

The first collection is called backpacks, and it just contains 5 backpacks. The only attribute is the backpack number.

backpacks

	a case		a case		a case		a case		a case
---	--------	---	--------	---	--------	---	--------	---	--------

The second collection is the Sample of backpacks. I needed to sample all 5 without replacement, so the cases are the 5 backpacks in the order chosen. There is a second attribute which I called person_num, but it's just the case index. I just want to see if the backpack number and the person number match. Then there is one measure which counts how many matches there are.



Sample of backpacks

backpacks

	bkpk_num	<new>
1	1	
2	2	
3	3	
4	4	
5	5	

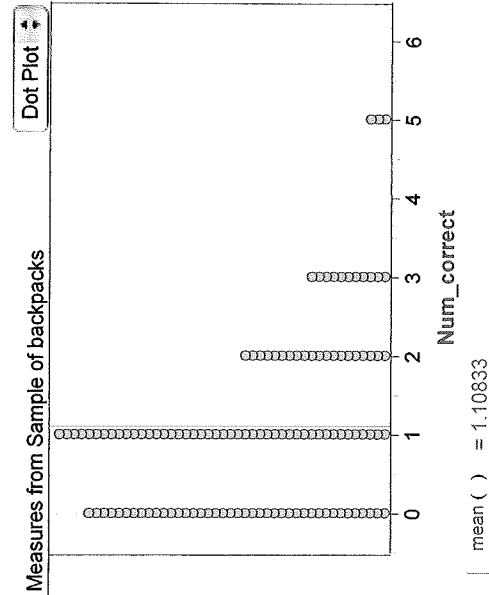
Sample of backpacks

	bkpk_num	person_num	<new>
1	3	1	
2	2	2	
3	1	3	
4	4	4	
5	5	5	

The third step of the simulation is to collect measures, so this collection is from Sample of backpacks. I have it set to collect 120 measures, since the theoretical answer is $1/120$. A case here is a set of 120 trials of 5 backpacks. The single attribute is the number of people who got their own backpack. What I am interested in is the number of measures for which the number correct is 5. I have plotted the mean number correct, so I can see this number also. After trying this several times, the mean seems close to 1.



Measures from Sample of backpacks



In order to really use the simulation to try to answer the question, I have created a fourth layer of analysis. First I added a measure to the previous collection, called `five_correct`, which counts the number of times you get 5 correct in 120 samples. I expect the mean of this number over a large number of trials to be 1, since the theoretical probability is $1/120$, and in the example shown, it's quite close. It's currently set to collect 100 measures at a time (since this is 100 times 120, it takes a LONG time). Since every step is set to replace existing measures, if you want to re-run the whole simulation, you just click Collect more Measures in this collection.

I feel that adding this step to the simulation helps in terms of answering the question posed. However, Measures from Measures from Samples seems like a complex structure and perhaps hard to follow, so I'm feeling ambivalent about whether it helps or hurts overall.

